DOCUMENT RESUME

ED 447 604 EA 030 779

AUTHOR	Rouk, Ullik
TITLE	"Tough Love": State Accountability Policies Push Student Achievement.
INSTITUTION	Southwest Educational Development Lab., Austin, TX.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	2000-08-00
NOTE	18p.
CONTRACT	RJ96006801
AVAILABLE FROM	Southwest Educational Development Laboratory, 211 East Seventh Street, Austin, TX 78701. Tel: 512-476-6861; Web site: http://www.sedl.org.
PUB TYPE	Collected Works - Serials (022)
JOURNAL CIT	Insights on Education iPolicy, Practice, and Research; n11 Aug 2000
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Achievement; Academic Standards; Access to Information; *Accountability; Achievement Gains; Achievement Tests; Disability Discrimination; *Educational Assessment; Educational Testing; Elementary Secondary Education; Government School Relationship; Incentives; *Parent Attitudes; Program Effectiveness; Sanctions; School Choice; *School Effectiveness; School Publications; *Student Evaluation; Test Results; Test Reviews; Testing
IDENTIFIERS	Individuals with Disabilities Education Act; *Southwest Educational Development Laboratory

ABSTRACT

State leaders are using complex accountability systems, composed of standards, assessments, public reporting, rewards, and sanctions, to raise student, school, and district achievement. The public strongly supports making academic standards more challenging, despite a lack of consensus on content and outcomes. Some states may revise standards until other reforms are in place. Student scores are now the primary indicator of district, school, teacher, and student achievement, with 48 states administering statewide testing using a mix of tools, including norm-referenced tests, criterion-referenced tests, performance assessments, and some evaluating attendance and dropout rates. The significant consequences of testing raise concerns for some parents, civil-rights activists, and educators. State policymakers must make decisions carefully, determining educational value, and working to gain public support on test design and use. States bear the expense of developing and carrying out testing, and must decide whether or not to control for prior achievement, family, and community characteristics. Public reporting helps the public understanding, and builds and sustains support for accountability systems. Union opposition often complicates use of rewards and sanctions. Evidence of limited success does exist for reconstitution. States must decide whether, and how, to include special-needs students, students with disabilities, and English learners in assessment systems designed to promote continuous improvement. A comprehensive system incorporates professional development, high standards, and student assessment, and many states now recognize the expense and effort this requires. Some states use safeguards to prevent

Reproductions supplied by EDRS are the best that can be made
from the original document.

testing manipulation. States may increasingly guide development of accountability systems that use student performance to begin discussions, link performance with classroom practice, and focus on improving education for all students. (Contains 22 references.) (TEJ)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 417 601

Insights
ON EDUCATION POLICY, PRACTICE, AND RESEARCH

AUGUST 2000
NUMBER 11
SOUTHWEST EDUCATIONAL
DEVELOPMENT LABORATORY

"Tough Love:"

State Accountability Policies Push Student Achievement

States have assumed a broader role in education than ever before. Across SEDL's region and others, state leaders are promoting systemic improvement with tough accountability policies.

Education in the United States has always been a state responsibility. But until recently, most states limited their activities to licensing teachers, accrediting schools, and issuing rules and regulations on matters such as the length of school day and year. State education agencies in a small number of states, Louisiana and Texas among them, also adopted textbooks.

Today, all that has changed. State leaders, including those in Arkansas, Louisiana, New Mexico, Oklahoma, and Texas, are leveraging improvement through complex accountability systems made up of standards, assessments, public reporting, rewards and sanctions—all designed to raise student, school, and district achievement.

Some states have produced impressive results, winning wide public support for groundbreaking accountability procedures. Texas, for instance, has gained national recognition for its insistence on holding schools accountable for helping poor and minority students meet the same achievement levels as other students. Its efforts have been rewarded with not only higher test scores, but also with the ability to remove some schools from the ranks of "low performing."

Others are moving more cautiously or, like Oklahoma, are still wrestling with ways to finance their accountability systems. And a growing number are finding themselves up against the political ramifications of their new measures and contemplating mid-course corrections. Among the issues these states are reexamining are: Should states and districts hold schools to absolute standards or adjust their expectations to account for differences in family income or existing student achieve-

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CE030174

BEST COPY AVAILABLE

State leaders are leveraging improvement through complex accountability systems all designed to raise student, school, and district achievement.

ment? How fast should scores improve? What if too many students fail?

This issue of SEDL's *Insights on Education Policy, Practice, and Research* reviews some of the assumptions, tensions, and lessons in the current accountability movement as well as the progress states in the region have made in implementing their accountability policies. From time to time, it also draws on the experiences from other states across the country.

ELEMENTS OF STATE ACCOUNTABILITY SYSTEMS

The words assessment and accountability are often used synonymously. Accountability, however, is a much broader concept in which assessment is but one element. The other elements in these new systems are standards, public reporting, rewards and sanctions, and continuous improvement. Those held accountable include districts, schools, teachers, and students.

Standards

The purpose of standards is to provide teachers and students with clear expectations for instruction and learning. They signal that all students, regardless of family background or where they happen to live, are expected to achieve a high level of learning.

Accountability systems contain two kinds of standards. States and communities have established standards for the content of education in different grades—that is, what students need to learn and know—and performance standards—the level of acceptable proficiency within grades.

According to *Quality Counts*, a state-by-state review of education reform published by *Education Week*, every state but Iowa has adopted standards for some subjects. Of the 50 states, 44 have standards in English, mathematics, social studies, and science. (The sidebar on page 3 provides information on the status of the development and adoption of standards in the SEDL region.)

Higher Standards Win Public Support

Public opinion remains staunchly behind higher, more challenging standards. More than 80 percent of parents polled by Public Agenda (Johnson & Duffett, 1999) said they believed higher standards would strengthen students' academic performance. What's more, support for higher standards was strongest among low-income families—historically those most poorly served by the education system. Even students themselves said that they would work harder if they were more challenged to do so.

But the adoption of standards assumes society's consensus on both the academic content and performance outcomes of schooling. And there's the rub. Despite the ubiquity of education standards, the dispute over how to mesh these new expectations with the everyday realities of schooling has never been resolved.

How High Is Too High?

One consequence of higher standards has been low student test scores. Low scores on state mandated tests are setting off policy debates about "How high is too high?" Massachusetts, New York, and Virginia are a few of the states exploring options for dealing with poor performance. Such options

BEST COPY AVAILABLE

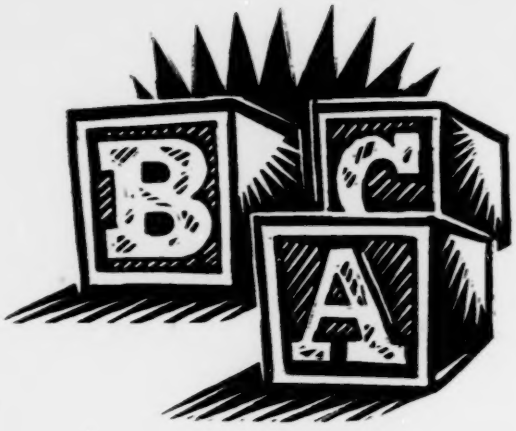
include lowering the passing bar or postponing consequences for schools and students until other parts of the system, such as professional development and mechanisms for providing extra help to low-achieving students, are in place.

In his State of American Education address in February, U.S. Secretary of Education Richard W. Riley reminded listeners, "Setting high expectations does not mean setting them so high that they are unreasonable except for a very few."

But other critics of the education system argue that softening expectations or reining in the implementation of higher standards now would signal a wavering resolve among policymakers to press reform forward. There are those, too, who believe that the rigor, depth, and scope of student standards still lag behind the knowledge and skills that students will need for success later in life. These people say the new standards are still not high enough. Questions about the rigor and usefulness of the new standards led to the formation of Achieve, Inc., a nonprofit group based in Cambridge, Massachusetts. Created by the nation's governors and business leaders, Achieve, Inc. has as its mission helping states strengthen the quality of their standards and tests by benchmarking them against those of other states and nations.

Assessment

Achieving high standards is the essence of accountability. To measure how well schools and students are meeting high standards, states have developed new assessment systems, or are refining existing systems, to align with the standards. Student scores on these assessments have become the number one indicator of district, school,



Standards in the Region

In Arkansas. The state has completed content standards for mathematics, language arts, science, reading, history/social science, foreign language, music, and health/physical education. It has developed performance standards for mathematics and literacy for grades 11 and 12 and is now developing them for grades 4 and 8.

In Louisiana. The state has produced content and performance standards in mathematics, language arts, social studies, science, the arts, and foreign languages.

In New Mexico. The state adopted content standards in the arts, languages (modern, foreign and native), mathematics, science, social studies, physical education, health education, and career readiness. It developed performance standards in each area, which are being reviewed and in some cases, field tested.

In Oklahoma. The state has written content standards for language arts, mathematics, science, social studies, languages (including foreign, Native American, and American Sign), the arts, technology, health, safety, physical education, and information skills. Performance standards have been determined for grades 5, 8, and 11 for writing, reading, mathematics, science, U.S. history and government, geography, the arts, and Oklahoma history (grade 11 only).

In Texas. The state has standards in language arts/reading, mathematics, science, and social studies. Performance standards are incorporated into each content area.

Source: Council of Chief State School Officers. (1999). *Status Report: State Systemic Education Improvements 1999*. Washington, DC: Author.

BEST COPY AVAILABLE

States have come to understand that their use of a test must match the purpose for which it was designed.

teacher, and student achievement.

Most states use a mix of tools to measure student performance, including norm-referenced tests, criterion-referenced tests, and performance assessments. They seek a balance, combining open-ended formats that ask students to "invent" solutions to problems with more traditional standardized, norm-referenced tests. At one time, many states dropped multiple-choice test items in favor of performance and portfolio assessments. Testing experts believe these assessments provide a more accurate picture of what students know and can do, but questions about the reliability of such tests have since caused some states, including Kentucky a leader in education accountability, to reintroduce multiple choice (Whitford & Jones, 1999). Others put some multiple choice back into assessments to reduce cost and time requirements.

According to *Quality Counts*, 48 states are administering statewide testing programs, and 37 said they incorporate "performance tasks" in their assessments. Among the 48, 41 have aligned their tests in at least one subject to standards. *Quality Counts* reports that 21 states have aligned their standards and tests in all four primary academic subjects.

In addition to using student test scores, states often gauge how well schools and districts are doing by looking at factors such as attendance and dropout rates. In Louisiana, for example, high school students' scores on the state test account for 60 percent of a school's score and scores on the national test account for 30 percent. The remaining 10 percent is determined by a school's student attendance and dropout rates.

Controversy over High Stakes

No one disputes that testing has a place in state accountability systems. Yet, the nature of these tests has steeped them in controversy. Many tests are "high-stakes," meaning that they have significant consequences for students and schools that do not meet achievement expectations. For example, "high stakes" come into play when students are denied promotion or high school graduation because of their low performance on tests, or when a school is totally reorganized because of recurrent low test scores.

Advocates of testing insist that the objectivity of test results ends the uncertainty about what students know and don't know. Local districts and schools can use test results to identify their instructional strengths and weaknesses, and make decisions about their instructional programs. Critics warn that high-stakes tests can distort and narrow the purpose of schooling to the quest for test scores (WestEd, 2000). High-stakes tests encourage teachers to focus solely on what is tested, obscure richer ways of judging schools, and place blame for ineffective teaching on students. Rather than using test scores to judge students and schools, some assesment experts recommend using test scores as one among many sources of information to answer the same questions about students and schools. They also argue that testing instruments and technology are not up to the demands that high-stakes accountability places upon them (Linn, 2000).

Even as high-stakes testing becomes integrated into the system, many parents, civil rights activists, and educators are questioning the wisdom of relying on test scores for such decisions as student promotion and high school graduation. Parents in one of Michigan's most affluent school districts recently rebelled against a new high school proficiency test that they

BEST COPY AVAILABLE

claimed did nothing but embarrass students bound for college. Arguing against the inflated value of one test and the loss of local control, they organized student boycotts, political lobbying, and lawsuits to resist the test. Such tensions are making policymakers listen and sometimes change their plans.

Tough Decisions for Policymakers

Putting state assessment programs into place is filled with tough decisions, each one creating its own tensions for policymakers. Decisions have to be weighed carefully, both to ascertain their educational value and to gain public support around issues of the design and appropriate use of tests.

Most state decisionmakers have learned that no single measurement instrument can do all things well. Tests designed to hold schools publicly accountable for student achievement are not the same tests that identify weaknesses or guide instruction; neither can they be used to set improvement targets for schools and districts. States have come to understand that their use of a test must match the purpose for which it was designed. Consequently, they've had to decide what they want their assessment programs to do, and develop—or select—a range of assessment strategies accordingly. States have had to decide whether to develop their own assessments, designed specifically to address their own standards, or to rely on commercial assessments. Developing new tests that are aligned with standards is a major expense for states. States often have few resources available for this development. Purchasing tests may cost considerably less. However, while test publishers do try to align their test items with common elements in state standards, these

Assessment in the Region

In Arkansas. Over the next four years, Arkansas will phase in the Arkansas Comprehensive Testing, Assessment, and Accountability Program. The plan incorporates standards, professional development for teachers, and state tests for students in grades 4, 6, and 8. Beginning in 2003-04, schools will be assessed according to state test scores, attendance, graduation rates, school safety, and teacher qualifications.

In Louisiana. Fourth and eighth graders in Louisiana took for the first time this spring LEAP 21 (Leap into the 21st Century) —a new state test designed to end the promotion of students who are not academically ready to enter the next grade. Districts must offer summer school programs to students who fail, after which students have another opportunity to take the exam. Students in tenth and eleventh grades will also take the test beginning in 2001. Last fall, K-8 schools were rated on the basis of student performance on the exam, as well as on the Iowa Test of Basic Skills, student attendance, and dropout rates.

In New Mexico. The accountability system will go into effect during the 2000-01 academic year. The assessment portion of the system consists of the CTBS/Terra Nova Exam in mathematics, science, language arts, and social studies for grades 3 through 9. Students are also administered a writing examination in grades 4 and 6 and a reading assessment in grades 1 and 2. Students must master the tenth-grade standards in order to graduate from high school but have until their senior year to pass the exam.

In Oklahoma. The statewide assessment program, Oklahoma's Core Curriculum Tests, assesses fifth and sixth graders in mathematics, reading, science, U.S. history, geography, arts, and writing. Eleventh graders are also tested in geography. The core curriculum tests for eleventh-grade students were discontinued in 1999, and will be replaced with "end-of-instruction tests" starting next school year.

In Texas. Every spring, Texas students in grades 3 through 8 and in grade 10 take the Texas Assessment of Academic Skills (TAAS) exams. Beginning in 2002-03, ninth and eleventh grade students will take the exams as well. TAAS scores combine with attendance and graduation rates to give schools a state accountability rating. Texas also requires that students pass examinations in reading, writing, and mathematics in order to graduate from high school. These exams, which test students on content through the end of ninth grade, are first administered in tenth grade. Students must pass all three exams but retake only that which they failed.

Sources: Arkansas Comprehensive Testing, Assessment & Accountability Program, July 1999; *LEAP for the 21st Century High Stakes Testing Policy*, Louisiana, May 2000; *Handbook: Statewide Student Assessment System. Information for Parents, Students, Teachers, and other School Personnel*, New Mexico, June 1999; communication with staff of State Department of Education and Office of Accountability, Oklahoma; *Quality Counts 2000, Education Week*.

BEST COPY AVAILABLE

Most state decisionmakers have learned that no single measurement instrument can do all things well.

tests are unlikely to align as closely as items in a test developed by the state.

States also have had to decide what to compare and how often to test students. Comparing one year's fourth-graders against another's may not provide a true picture of achievement because the test population is not the same. This was one of the most significant controversies in the implementation of Kentucky's accountability system. State officials responded by spreading testing to more grades (Whitford & Jones, 2000). Some experts recommend annual testing at each grade level, arguing that annual testing localizes student performance to the most natural unit of accountability, the grade level or classroom. It also yields the most up-to-date information and limits the amount of data that is lost when students move to other schools and districts. While measuring individual student progress each year offers a more accurate assessment, this method is expensive and difficult to carry out among highly mobile student populations.

At the same time, states have had to decide whether to measure absolute performance or growth in performance. Some states, like Arkansas, recognize schools for both absolute levels of achievement and for growth. Louisiana schools, on the other hand, are given a growth target to reach within two years. In making these decisions, states have had to decide what is an acceptable level of performance and what constitutes satisfactory progress. Other questions to address include the following: Should the same rate of progress be expected all the time? How much growth is reasonable to expect? Should the same amount of growth be expected from schools that start at different achievement levels?

Finally, states have had to face the particularly prickly issue of whether to control for differences in student, family, and community characteristics across students. Some districts believe that controlling for differences in prior achievement and student, family, and community characteristics across schools "institutionalizes low expectations for poor, minority, low-achieving students" (Elmore, Abelmann, & Fuhrman, 1996). Others argue that using data on these characteristics effectively would require collecting them for all students, increasing the data burden for districts, something only the largest districts may be prepared to handle. Most others generally have on hand only the limited administrative data that is available on students' race, gender, eligibility for free or reduced-price lunches, special education, or limited English proficient (LEP) status.

Public Reporting

Reporting school and district test scores to the public has become the major tool by which states demonstrate accountability. As of 1999, all states but one reported student achievement results to the public. Of those, 47 provided information at the district level and 41 reported their data at the school level (Council of Chief State School Officers, 1999), typically by using anywhere from four to six levels of proficiency in each subject area, ranging from the best to the worst. These school and district reports generally indicate the percentage of students who have reached each level of proficiency. Because test results remain the dominant means by which policymakers and the press describe the performance of schools, accurate reporting of student and school performance has taken on critical importance in state accountability systems.

BEST COPY AVAILABLE

A major challenge in producing such accountability information is explaining the data in ways that the public can understand (National Research Council, 1999). Student reports are available directly to students and their parents. In general, they provide information on the student's level of accomplishment in each content area tested and indicate how the student scored with respect to the statewide standards. The reports also describe typical performance at different standard levels so parents will know how to interpret the results.

Other community members interested in education generally get the news about test scores through the media. Good reporting about the quality of schools performs a valuable public service; inaccurate or incomplete reporting can add fuel to already seething controversy. Responsible reporting fully explains test scores and their limitations. It does not oversimplify or delete the very information that helps the public understand the meaning of scores. Nor does it make invalid statistical comparisons, for example, ranking schools that in reality had statistically negligible differences on their test scores.

Communications and public relations are of tremendous importance because the public's understanding of these technically complex accountability systems will ultimately affect the public's confidence in them. What is the pedagogical basis of a new assessment strategy? Why is one level of achievement proficient and another not? Why did some schools receive sanctions and others rewards? Why does this testing program need to be revised already? Helping the public understand such issues is groundwork that will sustain commitment and support for an accountability system

Public Reporting in the Region

In Arkansas. By November 15 of each year, every Arkansas school district must report, in a general circulation newspaper in the district, its progress toward accomplishing its program goals, accreditation standards, and proposals to correct deficiencies. Districts also issue school performance reports that are sent to parents by mail and disseminated to the public on the Department of Education's Web site. Each school is required, at least once a year, to explain its policies, programs, and goals to the community in a public meeting. Arkansas also analyzes and reports assessment data separately for special education students, students who are not proficient in English, and students who frequently move.

In Louisiana. The state reports annually on its progress in reaching its 10-year goals. It also publishes report cards for each school. These provide a school performance score, the school's progress in reaching its growth target, the school's performance compared to "like" schools, and subgroup performance data such as race, gender, and high-poverty vs. non-poverty student achievement. Data have been made available on the State Department of Education's Web site.

In New Mexico. Each year, the New Mexico State Department of Education publishes a State Accountability Report. This report contains information on education programs, trends, finances, and student achievement. In addition, the New Mexico legislature passed a key statute in 1997 requiring each school district to publish an annual School District Report Card. It provides school-based data on student achievement, attendance, dropout rate, school safety, and parent and community involvement.

In Oklahoma. The Oklahoma State Department of Education issues district report cards, a report of student test scores, and a reading sufficiency report. The Office of Accountability, also issues reports for the state, districts, and individual schools that describe community characteristics, district educational processes, and student performance. The reports are available on the Web. School-level report cards are also distributed to schools so principals can add their own comments and provide them to parents. District reports are distributed to libraries, a mailing list, and others who request them.

In Texas. The state's indicator system combines TAAS scores, end-of-course passing rates, attendance, high school completion rates, and other information to assign ratings to each school and district. The state has developed an extensive Web site to disseminate these data. Schools are responsible for disseminating a separate school report card to the parents of their students. The state also publishes district-level and state-level data and makes them available.

Sources: Arkansas Comprehensive Testing, Assessment & Accountability Program, July 1999; *LEAP for the 21st Century High Stakes Testing Policy*, May 2000; communication with Department of Education staff in New Mexico; Oklahoma Office of Accountability Web site, http://www.schoolreportcards.org/, Texas Education Agency Web site, http://www.teastate.tx.US/tea/account.html/

BEST COPY AVAILABLE

Rewards and Sanctions in the Region

In Arkansas. The state has plans to reward exceptional performance in two categories: performance, which recognizes absolute levels of student achievement, and growth and improvement, which recognizes upward trends. Awards will be phased in over time and may include cash awards that schools can use to expand programs, purchase materials, add technology, or give staff bonuses or other incentives. Any school that fails to achieve expected levels of student performance is listed as a "high-priority school." Schools that do not improve are placed on "alert status" the second year, "low-performing status" the third year, and "academic distress" the fourth year. These schools will be required to develop state-approved improvement plans and receive state technical assistance. The state may take over schools that fail to improve, although no timetable for such action has been established.

In Louisiana. Schools that meet their growth target and demonstrate improvement in the performance of students who are classified as high-poverty may receive monetary rewards. School personnel decide how the money is spent but cannot spend it on salary stipends. Schools that fail to meet growth targets are subject to three levels of corrective action. Initially, they are provided with assistance from District Assistance Teams in identifying their needs, developing school improvement plans, and examining their use of resources. Schools that do not improve after two years will be assigned a highly trained Distinguished Educator to work with them. At this point, parents will have the right to transfer their child to a higher performing public school. At a third level of corrective action, which again takes place after another two years, districts must develop reconstitution plans for schools that do not show 40 percent growth toward their targets.

In Oklahoma. The state is developing a rewards system and redesigning its support to schools. Under the present system, schools that are low-performing—those whose students score in the bottom quarter on a state-developed test and below the national norm on the Iowa Tests of Basic Skills—receive technical assistance from state department staff if they do not improve after one year. Technical assistance becomes even more intensive for schools that are still low-performing in the third year. The state has closed two schools over the past 10 years for poor achievement and fiscal management.

In New Mexico. The state recognizes schools that have made the most progress from one year to the next. The top 10 percent of these schools receive awards ranging from $1,000 to $14,000. Schools scoring in the top 25 percent on state assessments receive a certificate of recognition. Low-performing schools, on the other hand, must implement improvement plans. A school has two years to demonstrate improvement, after which the state can impose a curriculum, take over the school, or close it. In the last two school years, 25 schools have been deemed low-performing.

In Texas. Schools qualify for financial awards when they have: 45 percent of all students and of each student group passing each section of the TAAS, a dropout rate of 6 percent or less, and an attendance rate of at least 94 percent. A school must also rank at the top 25 percent of its comparison group in reading and mathematics. The system features separate recognition for schools that increase the number of parents or guardians attending parent-teacher conferences. Districts that find themselves in need of outside support, intervention, and mediation receive help from a peer assistance team. The formation of this team is part of an initiative sponsored by the Texas Education Agency, Texas Association of School Boards, and Texas Association of School Administrators.

Sources: Arkansas Comprehensive Testing, Assessment & Accountability Program, July 1999; LEAP for the 21st Century High Stakes Testing Policy, May 2000; *Quality Counts 2000:* Texas Education Agency Web site.

when some students start to fail or when schools find themselves labeled "unsatisfactory."

Rewards, Sanctions, and Assistance

The consequences of these new get-tough accountability policies are significant, both for educators and for students. Without consequences, so the thinking goes, educators and students will have little incentive to change.

For educators, the consequences of student test scores usually come in the form of rewards or sanctions. Thirteen states provide awards to schools that have met their achievement goals. Such awards are frequently in the form of cash that schools may use to expand their programs, purchase additional instructional materials or equipment, including technology, or give staff bonuses. Schools that do not meet their achievement goals may face sanctions ranging from having to obtain state approval for their improvement plan, having their accreditation revoked, or having to reorganize the entire school and staff. Eighteen states levy school-wide sanctions while several others have the authority to pose special measures such as allowing students to attend other public, and sometimes private, schools (*Quality Counts*, 2000).

Student rewards can take the form of an afternoon pizza party or a field trip. Sanctions can range from having to be tutored, attending summer school, being held back a grade, or not getting a high school diploma.

Combining incentives for performance with interventions and consequences for failure is typical in state accountability systems. A number of factors, however, make these practices especially thorny issues. For one, accountability systems have also become chief mechanisms for evaluating teachers, principals, and other administrators. This practice has drawn fire from teachers' unions, which argue that student test scores reflect more factors than those under a teacher's direct control; therefore, they have no place in personnel evaluations.

Also, the research on the complex nature of rewards and sanctions and its implications for education accountability is limited at best (Cohen, 1996). Researchers know little about what rewards and penalties might lead students to high levels of learning, the kinds of rewards that are effective in encouraging teachers to change their instructional techniques, or even whether rewards actually motivate students and educators to produce more.

The little that is known points out that rewards and penalties should be neither too trivial nor too heavy handed (Cohen, 1996). Rewards that recognize teachers' intrinsic motivation by, for example, sending them to workshops where they can add to their skills and knowledge are often more relevant, effective, and appreciated than outright cash bonuses.

The last step in overhauling schools that chronically fail to meet their performance targets is known as reconstitution. Reconstitution can occur when low-achieving schools, despite technical assistance and additional professional development, repeatedly fail to meet performance expectations. The practice essentially consists of ousting a school's teachers and administrators—and sometimes support personnel as well—and starting over from scratch, or at the very least, asking staff to reapply for their jobs. The goal is to replace a flagging school culture

Accurate reporting of student and school performance has taken on critical importance in state accountability systems. A major challenge in producing such accountability information is explaining the data in ways that the public can understand.

BEST COPY AVAILABLE

What little is known about incentives for teaching points out that rewards and penalties should be neither too trivial nor too heavy handed.

with one that supports high standards.

Schools in SEDL's region have escaped such dramatic shakeups by their states. However, states are not alone in their authority to reconstitute schools that fail. Districts can reconstitute schools as well, as was the case in San Antonio. Schools in Cleveland, Chicago, Denver, Prince George's County (Maryland), and San Francisco have also been reconstituted. Where entire school districts have been reconstituted, as happened in three New Jersey districts, the state removed local school board members, the superintendents, and other key personnel. The state appointed policymakers for each district until new school boards were created.

The effectiveness of such housecleaning has not been fully proven, but limited evidence of success does exist in some locales. After the state took control of schools in Cleveland, the district's state proficiency scores showed a slight upward trend. Data from Paterson, New Jersey, indicate student gains in reading, writing, and mathematics. A middle school in San Francisco demonstrated exceptional improvement on its standardized test scores.

Advocates argue that the threat of reconstitution serves as an incentive to keep educators' eyes on student performance targets. Not surprisingly, critics of reconstitution, teachers unions foremost among them, claim that reconstitution is a simplistic response to a complicated problem. Sandra Feldman, president of the American Federation of Teachers, characterized reconstitution as "getting rid of people instead of bad practices" (Hendrie, 1997). Others claim that the changes wrought by reconstitution are largely cosmetic because displaced teachers typically resume their careers in other schools.

Despite union resistance, however, reconstitution appears here to stay, at least for the immediate future, if a clause in Detroit's contract with teachers is any indicator. The deal between the district and teachers there allows reconstitution of schools that have lost state accreditation and failed to improve despite extra help. Union leaders in other districts such as Chicago and Cincinnati have negotiated interim steps between putting schools on probation and reconstituting them, thus giving schools more time to improve.

WHAT ISSUES DO STATES FACE IN PUTTING THEIR ACCOUNTABILITY SYSTEMS INTO EFFECT?

New challenges and needs face states as they press to put their accountability systems into effect.

Equity Issues

Historically, states and districts have excluded or waived participation of students with special needs from assessment mandates. New laws, however, require their full inclusion. Including special needs students, according to proponents, provides more accurate assessment results and signals a commitment on the part of states and districts to support the academic progress of all students. At issue, however, is how to accommodate special needs students without overcompensating (Shepard, Taylor, & Betebenner, 1998). With little data and few models for guidance, the challenge for states has been considerable.

Some states in the SEDL region

BEST COPY AVAILABLE

such as New Mexico and Texas, have achieved national recognition for their groundbreaking work in developing and implementing alternate or modified exams for Spanish-speaking and special education students. Other states are struggling to develop strategic plans for improving inclusion practices in all aspects of assessment and accountability, not just for legal and procedural compliance, but to meet broader educational equity goals.

Students with Disabilities

To be eligible for funding under Part B of the Individuals with Disabilities Education Act (IDEA), 1997 amendments require states and districts to include children with disabilities in assessment programs. Furthermore, the law stipulates that states and districts must make appropriate accommodations and modifications in administration, if necessary, and provide alternate assessments for children who cannot participate in the general assessment program. Reports to the public should be made available with the same frequency and in the same detail as reports on the assessment results of non-disabled children. Furthermore, the data in these reports should take two forms: aggregated to compare the performance of children with disabilities with all children, and disaggregated to compare the performance among all children with disabilities.

States have developed numerous accommodations for students with disabilities. Louisiana, for example, allows a limited number of special education students to participate in out-of-level testing or alternate assessment in lieu of taking general statewide assessments. Specifically, a maximum of 1.5 percent of students at any grade level per school district may participate in alternate assessment as determined by the LEAP Alternate Assessment Participation Criteria. Up to 4 percent of students at any grade level may participate in out-of-level testing. These students need to be functioning at least three grade levels below in reading, language, or mathematics. In the 1999-2000 school year, school districts were granted authority to waive the state's grade promotion policy for students with disabilities.

English-Language Learners

Similarly, states are finding ways to accommodate students' English language skills. The most typical accommodation is to allow English language learners to demonstrate their skills and knowledge in ways not hindered by language. This may involve translating assessments into students' native languages (a costly procedure that is not always feasible given the number of languages spoken in many schools in the United States). Other approaches include oral reading of the test in English, allowing students to use dictionaries, extending testing time, or changing the method of response from paper and pencil to performance.

Texas will have tests in Spanish in school year 1999-2000. In addition, Texas holds schools accountable for showing significant growth in the scores of sub-populations of students. This indicator focuses educators' efforts on each ethnic group and assures girls and boys are treated equitably.

Despite these accommodations, there is wide agreement on the need to develop better ways of assessing students whose lack of English language skills may misrepresent their achievement. Until this issue is resolved, says a report of the White

Despite these accommodations, there is wide agreement on the need to develop better ways of assessing students whose lack of English language skills may misrepresent their achievement.

BEST COPY AVAILABLE

Continuous improvement should be embedded in the language of every state accountability system. Without changing the ways schools operate, improvements are likely to be short lived.

House Initiative on Educational Excellence for Hispanic Americans, test scores should not keep Hispanic children from being promoted to the next grade or from receiving a high school diploma. Members of this commission instead proposed that test scores be used to hold schools accountable for providing an adequate education to Hispanic students (White House Initiative on Educational Excellence for Hispanic Americans, 1999).

Continuous Improvement

Continuous improvement should be embedded in the language of every state accountability system. Without changing the ways schools operate, improvements are likely to be short-lived. The very nature of current accountability policies demands that education systems implement new programs and teaching techniques, and that teachers have the expertise to teach to high standards. It goes without saying that teachers who have not yet developed that expertise require extensive professional development.

The real question about accountability systems may be: How can states foster true capacity building? In most cases, states have limited capacity to help local districts or schools grapple with issues of continuous improvement and professional development. States can provide the mandate. That is, the language of accountability systems can focus on those elements most critical to ongoing improvement. For example, systems show the importance of continuous improvement by awarding evidence of growth and improvement.

Most accountability systems of the states in SEDL's region describe indicators of gains in student achievement. In the development of Arkansas's system, for example, it became evident that teachers need support to understand what standards mean for instruction and assessment. As a result, professional development, along with high standards and student assessment, is one of the three components of its comprehensive system of accountability and assessment.

Access to data is another critical element of a state's accountability system. Being able to review data and make decisions based on those data is necessary if the staff of a school is to monitor and improve its performance. States that report data and make them available in multiple ways make it easier for a district or school to develop the ability to continuously check on its progress toward meeting its improvement goals.

Before teachers and principals can be held accountable for new and more effective programs, instructional techniques, and curricula, they must see a need for change and be willing to do things differently. Then they must receive professional development and continuous support to become comfortable with the new practices. Finally, they must be given time to demonstrate effectiveness.

The issue of teacher quality took center stage among governors, corporate CEOs, and education leaders during the third National Education Summit in Palisades, N.Y. This group put forth three recommendations for strengthening the teacher workforce. It proposed that:

- universities strengthen their teacher preparation programs to provide educators with the content knowledge and skills needed to help students meet higher academic standards;

- states create alternative pathways

BEST COPY AVAILABLE

into the teaching profession to attract the most talented candidates;

- professional development for teachers already in the classroom become a priority and emphasize tying such programs directly to standards.

Teacher quality and professional development continue to occupy policymakers' attention and are likely to be legislative issues in all five states that SEDL serves.

States' Capacity to Deliver

States place a great deal of emphasis on state-level assistance and intervention in schools or districts that aren't able to meet standards. Today, 19 states mandate that low-performing schools receive state assistance. States' actual capacity to deliver the personnel, expertise, and funding that these schools need to improve their practices, however, often can and does fall short (Elmore, Abelmann, & Fuhrman, 1996).

States' new, more systemic responsibilities can be challenged by the costs required to meet their accountability goals. Although some states claim to be leaner and smarter—Texas Education Agency (TEA), for example cut its personnel by 14 percent between 1980 and 1998—change is a human-intensive activity. To provide the technical assistance it promised schools, TEA contracted with 20 educational service centers, which operate on a combination of federal, state, and local money.

The situation is similar in other states. Sanctions and rewards don't come cheaply. Kentucky reports spending about $110,000 a year to assign experienced educators to work closely with schools designated "in crisis." At the other end of the spectrum, since 1995, it has spent more than $54 million in rewards directly to teachers and administrators (Fuhrman, 1999).

Some school officials now acknowledge, however, that they may have been asking for too much too soon. Accountability systems carry a high price tag. Hampered by the lack of funds, many state and local officials have been unable to put into place the technical assistance staff, professional development programs for teachers, extra help for students, and the other support necessary to meet suddenly accelerated standards. In some states, just getting the data to school districts in time for the information to be useful in planning has been a problem. In others, sufficient resources with which to train principals and teachers to analyze student performance data just aren't there. Before states can hold students, teachers, and principals accountable, they must have the capacity to deliver their own part.

Safeguarding Accountability Systems

One consequence of the added pressure on school administrators and teachers to put forth their best performance on school tests has been the manipulation of accountability systems. The media have reported instances of blatant cheating by teachers and administrators, but documented cases remain few.

Nonetheless, states have created new safeguards to protect both their test items and individual student test scores, going so far as to make student record tampering by school system employees a criminal offense. In most states, security is so important that testing officials go to great lengths to protect the integrity of

The real question about accountability systems may be: How can states foster true capacity building?

BEST COPY AVAILABLE

their systems. Many are holding workshops on test security for teachers and other school employees. In May 1999, the Texas state legislature passed a bill that makes it a felony to alter tests or test results. Those found guilty could spend up to 10 years in prison and pay up to $10,000 in fines. The action came on the heels of allegations that school employees in several districts had altered student information, including scores on the state assessment, to boost school ratings. Several teachers and staff resigned as a result and principals were severely reprimanded. Similar charges of test tampering have been levied against educators in districts around the country.

STATES' CONTINUING ROLE

Now that states have placed themselves in the center of school reform issues, they are unlikely to step aside anytime soon. Evidence on the importance of the elements in state accountability systems for student achievement became clear in a study of North Carolina and Texas (Grissmer & Flanagan, 1998). The researchers who conducted the study tied achievement gains in these two states directly to the state system of academic standards, assessments linked to standards, consequences for results, and other elements of the state infrastructure.

Given these findings, the role of states in guiding accountability may become even more significant. Sherman Dorn (1998) of the University of South Florida identifies three requirements that are common to meeting the need for accountability:

- Accountability systems should use student performance as a starting point for deeper discussion of educational problems.
- Accountability systems should link student performance with classroom practice.
- Accountability systems should focus on improving education for all children, not encourage schools to isolate and devote fewer resources to children who already have the odds stacked against them.

Accomplishing this new orientation will be an ongoing task. Policymakers will have to continually monitor the effects of their policies and modify them as they learn from their own experiences and those of others. Or, put another way, they will need to take a longer-term and more comprehensive perspective on accountability.

REFERENCES

Arkansas State Department of Education. (1999, June). *Arkansas Comprehensive Testing, Assessment, and Accountability Program*, Little Rock, AR: Author.

Cohen, D.K. (1996). Rewarding teachers for student performance. In S.H. Fuhrman and J.A. O'Day (Eds.), *Rewards and Reform: Creating Education Incentives that Work* (pp. 60-112). San Francisco: Jossey-Bass Publishers.

Council of Chief State School Officers. (1999). *State Education Accountability Reports and Indicator Reports: Status of Reports Across the States 1999*. Washington, DC: Author.

BEST COPY AVAILABLE

Dorn, S. (1998, January 2). The political legacy of school accountability systems. *Education Policy Analysis Archives, 6*. [On-line serial]. Available: http://olam.ed.asu.edu/epaa

Education Week. (2000, January 13). *Quality Counts 2000*. Author.

Elmore, R.F., Abelmann, C.H., & Fuhrman, S.H. (1996). The new accountability in state education reform: From process to performance. In H.F. Ladd (Ed.), *Holding Schools Accountable* (pp. 65-98). Washington, DC: The Brookings Institution.

Fuhrman, S.H. (1999, January). The new accountability. *CPRE Policy Briefs*. Philadelphia: University of Pennsylvania.

Grissmer, D. & Flanagan, A. (1998). *Exploring Rapid Achievement Gains in North Carolina and Texas*. Washington, DC: National Education Goals Panel.

Hendrie, C. (1997, September 24). Reconstitution gaining new momentum. *Education Week*.

Incentives for School Improvement Act (22-13A-5, NMSA 1978).

Johnson J. & Duffett, A. (1999). *Standards and accountability: Where the public stands*. New York: Public Agenda.

Linn, R.L. (2000). Assessments and accountability. *Educational Researcher, 29* (2), 4-16.

Louisiana State Department of Education. (1999, October 21). *Frequently Asked Questions About Out-of-Level Testing*. Baton Rouge: Author.

Louisiana State Department of Education. (Revised May 2000). *LEAP for the 21st Century High-Stakes Testing Policy*. Baton Rouge: Author.

Louisiana State Department of Education. (2000, July 1). *Louisiana's IEP Handbook for Students with Disabilities*. Baton Rouge: Author.

National Research Council. (1999). *Testing, Teaching, and Learning: A Guide for States and School Districts*. R.F. Elmore & R. Rothman (Eds.). Washington, DC: National Academy Press.

New Mexico State Department of Education. (1999, June). *Handbook: Statewide Student Assessment System. Information for Parents, Students, Teachers, and other School Personnel*. Santa Fe: Author.

Olson, Lynn. (1998, February 11). The push for accountability gathers steam. *Education Week* [On-line]. Available: www.edweek.org/ew/vol-17/22accoun.h17

Shepard, L., Taylor, G., & Betebenner, D. (1998). *Inclusion of limited English-proficient students in Rhode Island's grade 4 mathematics performance assessment*. (Report No. 486). Los Angeles: Center for the Study of Evaluation.

WestEd. (2000, February). The high stakes of high-stakes testing. *Policy Brief*. San Francisco: Author.

White House Initiative on Educational Excellence for Hispanic Americans. (1999). *Educational standards, assessment, and accountability: A new civil rights frontier*. Washington, DC: Author.

Whitford, B.L. & Jones, K. (1999). Kentucky lesson: How high-stakes school accountability undermines a performance-based curriculum vision. In B.L. Whitford and K. Jones (Eds.), *Accountability, Assessment, and Teacher Commitment: Lessons from Kentucky's Reform Efforts* (pp. 9-24). Albany: State University of New York Press.

This edition of *Insights on Education Policy, Practice, and Research* was written by Üllik Rouk, Policy Studies Associates in Washington, DC, and edited by Joyce Pollard, Director of Institutional Communications and program manager for policy information services at SEDL. Julia Guzman, SEDL policy specialist, conducted additional research.

Insights on Education Policy, Practice, and Research may be reproduced and copies distributed to members by educational agencies, organizations, and associations.

This publication is based on work sponsored wholly, or in part, by the Office of Educational Research and Improvement, U.S. Department of Education, under Contract #RJ96006801. The contents of this publication do not necessarily reflect the views of OERI, the Department, or any other agency of the U.S. Government.

SEDL | | OERI

Southwest Educational Development Laboratory
Wesley A. Hoover, Ph.D., President and CEO
211 East Seventh Street
Austin, Texas 78701
(512)476-6861
www.sedl.org

BEST COPY AVAILABLE

INSIGHTS AT A GLANCE

This edition of *Insights on Education Policy, Practice, and Research* examines the issues that states in SEDL's region have been addressing as they guide the development of complex accountability systems designed to improve student, school, and district achievement. This *Insights* also discusses the expansion of states' roles as comprehensive, statewide accountability systems mature.

The state's role in public education has ranged from setting the funding formula to regulating textbook approval and prescribing curriculum guidelines, to monitoring compliance with federal and state regulations. The development of statewide systems of accountability have created an expansion of that role to include: describing accountability as a system; striking balance among the components of the system—assessment and incentives is an example—to maintain the integrity of the system; and fostering capacity building for local districts.

Rather than a single concept, the idea of accountability as a system embraces an interconnected set of policies about:

- standards that provide teachers and students with clear expectations for instruction and learning,
- assessments that measure how well schools and students are meeting high standards,
- public reporting of student test scores, and
- rewards, sanctions, and assistance for teachers and schools.

Some states have produced impressive results, thereby gaining widespread support and national recognition. Others are refining their accountability systems in light of past experience. In almost every state, however, educators and policymakers are trying to resolve the tensions between setting new expectations and addressing unintended consequences as accountability systems play out in districts and schools. Among the issues states are grappling with are:

- How high should standards be?
- How fast should students improve?
- What mix of assessment tools to use?
- What indicators to measure and compare?
- How often to test students?
- Should assessments control for student, family, and community characteristics?
- How to report meaningful results to policymakers and the public?
- What are the effects of rewards and sanctions?
- How to accommodate disabled students and students who are not fluent in English?

In addition, continuous improvement should be the hallmark of a statewide accountability system. Unless districts and schools have the capacity to continuously monitor and respond to changing needs, any improvements they make are likely to be short-lived. Such capacity building involves:

- increasing teacher knowledge and skills,
- supporting principal leadership, and
- creating a professional community in which everyone has an eye toward high standards.

States, too, must have the capacity to deliver assistance to schools and districts. At present, states have little capacity to help locals deal with continuous improvement and professional development. Policies and practices related to state assistance and intervention can provide:

- a mandate for continuous improvement within the accountability system,
- alignment of parts of the system so they support continuous improvement,
- sufficient personnel to deliver assistance,
- expertise in data analysis and in linking student performance with classroom practice,
- funding, and
- safeguards that protect against the manipulation of accountability systems.

States are likely to remain at the center of school reform and improvement for a long time. Research shows that state accountability policies do influence academic standards and student achievement. Achieving a complete system of accountability, however, will remain work in progress. Policymakers will have to continually monitor the effects of their policies and modify them as they learn from their own experiences and those of others.

BEST COPY AVAILABLE